Research Report CCS 512 ✓

A STRUCTURE FOR CLASSIFYING AND
CHARACTERIZING EFFICIENCIES AND
INEFFICIENCIES IN DATA
ENVELOPMENT ANALYSIS

by

A. Charnes   W.W. Cooper
       R.M. Thrall*

CENTER FOR CYBERNETIC STUDIES

The University of Texas
Austin, Texas 78712

DTIC
JAN 2 2 1987

G

1 21 152

Research Report CCS 512

A STRUCTURE FOR CLASSIFYING AND CHARACTERIZING EFFICIENCIES AND INEFFICIENCIES IN DATA ENVELOPMENT ANALYSIS

by

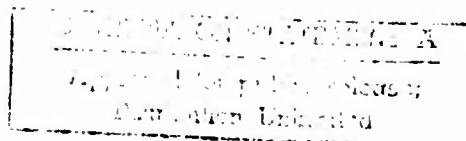A. Charnes   W.W. Cooper
R.M. Thrall*

Original Version: May 1985
First Revision: April 1986
Second Revision: September 1986

*Rice University, Houston, Texas

CENTER FOR CYBERNETIC STUDIES

A. Charnes, Director
College of Business Administration 5.202
The University of Texas at Austin
Austin, Texas 78712
(512) 471-1821

ABSTRACT

DEA (Data Envelopment Analysis) attempts to identify sources and estimate amounts of inefficiencies contained in the outputs and inputs generated by managed entities called DMUs (= Decision Making Units). Explicit formulation of underlying functional relations with specified parametric forms relating inputs to outputs is not required. An overall (scalar) measure of efficiency is also obtained for each DMU from the observed magnitudes of its multiple inputs and outputs without requiring use of a priori weights or relative value assumptions. A partition of DMUs into six classes is established via primal and dual representation theorems and three classification theorems which do not depend on non-archimedian analysis. Earlier theory is extended to explain the consequences of zero inputs and outputs and to utilize zero virtual multipliers (shadow prices). Three of the six classes are scale inefficient and two of the three scale efficient classes are also technically (= zero waste) efficient.

1. BACKGROUND AND INTRODUCTION.

Data Envelopment Analysis (DEA), as reported in [8], is an approach to measuring the efficiency of entities, e.g., not-for-profit entities, which have multiple outputs and multiple inputs that do not readily submit to treatment via unit prices or other a priori weights that might be used (e.g., in index numbers to secure comparability. In addition to providing meaningful scalar efficiency values, DEA is designed to identify sources and estimate the amounts of inefficiencies that might be present in the various output and input vectors. In these respects DEA may be likened to standard cost systems used in manufacturing enterprises although, unlike the latter, it does not require prior undertakings of extensive engineering studies. Also unlike the latter systems, it seeks to identify and estimate inefficiencies in outputs as well as inputs. Finally DEA seeks to accomplish all of this without requiring a priori specification of the underlying functional forms -- which can be a forbidding task in statistical regression analyses and like approaches to these same problems. See [6][1]

The basic theory for DEA as it exists at this point in time is set forth in [9] -- see also [8]. The purpose of the present paper is to relax some of the present conditions which restrict DEA uses to situations in which (a) all inputs and outputs are positive and (b) all variables are confined to positive ranges in the solution sets (see the discussion of D-proper virtual multipliers in Section 3). The principal results are: the primal and dual representations in Theorems 4 and 5, the

---

[1]See [13] and [14] for difficulties even when so-called flexible functional forms are used.

[2]We shall also abbreviate other references such as CCR [p. 41]

1

characterization of decision making units (DMU's) in Theorems 6, 7, and 8, and the analysis of the role of zero inputs and outputs in Theorem 18 and its corollaries.

Expanding the uses of DEA for evaluating possible additions or deletions of individual product lines is illustrative of possibilities for managerial uses at the level of individual Decision Making Units (DMUs). See [5] for a discussion of the possible interacting effects on efficiency evaluations that can occur with decisions to add or delete activities in the programs of a community college. The ability to bring DEA to bear on issues such as the evaluation of natural monopoly provides another class of examples in the arena of public policy analysis. See the treatment of "economies of scope" and "economies of scale" and their relation to issues of natural monopoly in the recent breakup of AT&T as discussed in [14] and [15].

To make this more precise assume that we have $j = 1,...,n$ DMUs, for each of which an observed vector of outputs y, and an observed vector of inputs x is available. Thus for DMU_j we have

(1)
$$y_j = \begin{bmatrix} y_{ij} \\ \vdots \\ y_{rj} \\ \vdots \\ y_{sj} \end{bmatrix}, \quad x_j = \begin{bmatrix} x_{ij} \\ \vdots \\ x_{ij} \\ \vdots \\ x_{mj} \end{bmatrix}$$

with $y_{rj} > 0$, $r = 1,...,s$ and $x_{ij} > 0$, $i = 1,...,m$ for every $j = 1,...,n$. Using the CCR ratio form from [11] we may then evaluate the efficiency of each (or all) of these DMUs relative to the others via

2

(2) $\max h_o (u,v) = \dfrac{u^T y_o}{v^T x_o}$

subject to

$$1 \geq \dfrac{u^T y_j}{v^T x_j} \, , \quad j = 1,\ldots,n$$

and

$$p < \varepsilon e^T \leq \dfrac{u^T}{v^T x_o} \, , \quad \dfrac{u^T}{v^T x_o}$$

where T indicates transposition. The vectors u and v are called "virtual multipliers", with s and m components, respectively, while the vector e^T has all of its elements to unity with the number of its elements determined by context. E.g., e^T, above is 1 x s and is multiplied by a "small" non-Archimedean constant, ε, by which we mean that each of the components of εe^T is greater than zero but smaller than any positive real number. In particular, given any positive real number z there is no real scalar k with the property

(3) $k\varepsilon \geq z$.

In the above formulation y_o and x_o are the output and input vectors for one of the $j = 1,\ldots,n$ DMUs incorporated in the constraints. Hence we have $\max h_o (u,v) = h_o^* \leq 1$, and via the following[1] we have:

---

[1]First given in [10] and reproduced in [8].

3

Non-Archimedean Efficiency Theorem:

Part 1: $h_o^* = 1$ if and only if DMU_o is efficient in all inputs and outputs.

Part 2: $h_o^* < 1$ if and only if DMU_o is inefficient in some inputs or outputs.

To obtain this result and also to place the DEA problem in readily computable form we used the Charnes-Cooper transformation of fractional programming, viz.,

$$(4) \qquad \nu = tu, \ \omega = tv, \ t \geq 0$$

to obtain the following dual pair of linear programming problems:

$$\max \ \nu^T y_o \qquad\qquad \min \ \theta \quad -\varepsilon e^T s^- \ -\varepsilon e^T s^+$$

subject to   subject to

$$(5) \qquad 1 = \omega^T x_o = tv^t x_o \qquad\qquad \phi = \theta x_o - X\lambda - s^-$$

$$\phi \geq \nu^T Y - \omega^T X \qquad\qquad y_o = \qquad Y\lambda \qquad - s^+$$

$$-\varepsilon e^T \geq \qquad -\omega^T$$

$$-\varepsilon e^T \geq -\nu^T \qquad\qquad \phi \leq \lambda, \ s^-, \ s^+$$

where the condition $\omega^T x_o = 1$ insures that the scalar t used in (4) will be positive so the reverse transformation may be used to secure (2) from (5).

All of the theorems of ordinary linear programming continue to hold for the non-Archimedean extensions included in (5).[1] Both problems have feasible solutions. Hence the dual theorem of linear programming gives the equality in

---

[1] See [8]

4

$$(6) \qquad \nu^{*T}y_o = \theta^* - \varepsilon e^T s^{-*} - \varepsilon e^T s^{+*} \leq 1$$

where the asterisks indicate optimum values and to establish the inequality one need only observe that $\omega^T X \geq \nu^T Y$ is required for the problem on the left in (5) and this requires $\omega^T x_o \geq \nu^T y_o$ which, together with the condition $\omega^T x_o = 1$, implies $1 \geq \nu y_o^T$ for feasibility.

The expression $\nu^{*T} y_o$ on the left of (6) can be a real number only if the components of s^{-*} and s^{+*} are all zero on the right. With $s^{+*} = 0$ it is not possible to improve the output performance by augmenting the corresponding y_o components by the amount of these non-zero slacks. But even if $s^{-*} = 0$ it may still be possible to improve the inputs, if the optimal value of the scalar $\theta^* < 1$. If, however, $\theta^* = 1$ and all slacks are zero then also $\nu^{*T} y_o = 1$ and all components of ν^{*T} are real valued. On the other hand if any components of s^{-*} or s^{+*} are not zero then we must have $\nu^{*T} y_o < 1$ and ν^{*T} must contain some non-Archimedean components. Indeed transposing back to (2) via (4) the sources (and amounts) of inefficiencies can be identified that produce $h_o^* < 1$ while, conversely, if $h_o^* = 1$ then also $\nu^{*T} y_o = 1$ and no inefficiencies are present.

Examples.

As these developments show, these non-Archimedean elements play an important role in the theory and possible uses of DEA. To explore possible consequences of eliminating these non-Archimedean elements we utilize the following example which involves $j = 1, 2$ DMUs each of which produces 1 unit

5

of output (indicated in the first row) and each of which utilizes two inputs in the amounts indicated in rows 2 and 3.

DMU

	1	2
y_{1j}	1	1
x_{1j}	1	1
x_{2j}	5	6

The DEA efficiency evaluation problems are, respectively,

for DMU_1

$$\max\ h(w) = \frac{u_1}{v_1 + 5v_2}$$

subject to

$$1 \geq \frac{u_1}{v_1 + 5v_2}$$

$$1 \geq \frac{u_1}{v_1 + 6v_2}$$

for DMU_2

$$\max\ h(w) = \frac{u_1}{v_1 + 6v_2}$$

subject to

$$1 \geq \frac{u_1}{v_1 + 5v_2}$$

$$1 \geq \frac{u_1}{v_1 + 6v_2}$$

All variables are merely constrained to be non-negative and not required to be positive as before. Here

$$w^T = (u_1,\ v_1,\ v_2)$$

is a member of the set of all "virtual multipliers" W when the elements of w are all non-zero and non-negative. If W_1 and W_2 represent the set of optimizing multipliers for, respectively, DMU_1 and DMU_2 in the obove example, we have $W_1 = W$ but $W_2 \neq W$ since an optimum is achieved for DMU_2 only when $v_2 = \phi$.

With $v_2 = \phi$ an optimum is achieved for DMU_2 with $h(w) = 1$ so that its behavior is incorrectly characterized as efficient in all inputs and outputs. This is a consequence of eliminating the non-Archimedean condition on the variables in (2), and, as already noted, the appearance of these non-Archimedean elements in an optimum solution make it possible to identify the sources of inefficiency and estimate their amounts by reference to the variables with which they are associated in (5).

In the section that follows we shall give an entirely new route which will associate with DMU_2 a multiplier set W_2 related to both sets of LP problems. In the above example we will find that dim $W_1 = 3$, which is the maximum number of postive components for variables in the above problem, whereas dim $W_2 = 2$ because $v_2^* = \phi$ in every optimum solution for DMU_2.

This suggests there may be a way of classifying DMUs by reference to the dimensions of their multiplier sets and that proceeding in this manner will enable us to retrieve what was lost by eliminating the non-Archimedean elements, ε. Thus, as was done in [16] we partition the set of all DMUs into the following six classes

$$E, E', F, NE, NE', NF$$

where the first three classes are on a frontier and the second three are not. The following figure adapted from p. 46 represents the situation for two inputs and one output. The solid lines connecting the points E and E' represent frontier points that are efficient. Points like F which are also on the frontier are not efficient since, as in the case shown, one of the

---

[1] That is, the DMUs associated with these points are at MPSS (Most Productive Scale Size) in the sense of Banker [2].

inputs may be reduced and this can continue until E is achieved without reducing output or increasing the other input.

All of the points E, E', F have optimal O values of unity in (5) and hence are scale efficient in the sense of Banker [2]. As already noted, however, F is not technically efficient. Finally the points NF, NE, and NE' are not on the frontier and hence are neither technically nor scale efficient.

In the above example DMU_1 is associated with E and DMU_2 with F since optimality is achieved with $h^* = 1$ in both cases. The sets E' and I are empty but, of course, this need not be true in other examples -- to which we shall turn after introducing suitable notations and definitions to cover all possible cases in the section that follows.

This paper is a technical supplement to [16] which will also be referenced as CCT. The other two principal references are [11], also referenced as CCR, which was the first of a long and continuous series of papers on DEA, and [8] which is a recent review and codification of DEA theory and contains an extensive bibliography for DEA.

This paper contains proofs of the new results announced in CCT [16]. Although we will rely for the most part on the notations and terminology of the three principal references the proofs require introduction of a number of new concepts and notations.

2. DATA DOMAINS.

Data Envelopment Analysis (DEA) is a framework for investigating the efficiency of a collection of Decision Making Units (DMUs) with common inputs and outputs see CCR [11], pp. 429-430 and [8], pp. 1-6.

8

We consider n DMUs, each with m inputs and s outputs. We define a data domain D as a (ordered) set

(7) $J = (DMU_1, \ldots, DMU_n)$

of DMUs and a matrix

(8) $P = [P_1, \ldots, P_m]$

with s + m rows and n columns. The jth column

(9) $P_j = \begin{bmatrix} y_j \\ -x_j \end{bmatrix}$

is made up from an input vector x_j whose ith component x_{ij} is the amount of input i used by DMU_j and an output vector y_j whose rth component y_{rj} is the amount of output r produced by DMU_j. We write

(10) $D = (J,P).$

We assume that the components x_{ij} and y_{rj} are positive or zero and require that for no j is either $x_j = 0$ or $y_j = 0$.

We say that a data domain D is in reduced form if for no pair (j,k) with $j \neq k$ and scalar α is

(11) $P_k = \alpha P_j.$

Let J^* be a subset of J, and let P^* be the submatrix of P consisting of those columns P_j for which $DMU_j \in J^*$. Then we call the data domain

(12) $D^* = (J^*, P^*)$

a subdomain of D.

Associated with a data domain D is a set of vectors

(13) $K(P) = \{p \mid p \leq P\lambda = P_1\lambda_1 + \ldots + P_n\lambda_n, \lambda \geq 0\}$

9

called the conical hull of P.

Let $D = (J,P)$ be a data domain and let D^* be the subdomain of D defined as follows. Let

(14) $J^* = $ (DMU$_j$ in N for which there is no $k < j$ with P_k proportional

to P_j).

Then, clearly, D^* is reduced.

3. VIRTUAL MULTIPLIERS AND DEA-SCALE-EFFICIENCY.

Let u and v be non-negative, non-zero vectors with s and m components, respectively. We call the vector

$$w = \begin{bmatrix} u \\ v \end{bmatrix}$$

a virtual multiplier and denote by W the set of all virtual multipliers. Next, for $w \in W$ and for any DMU$_j$ let $g_j(v) = v^T x_j$, $f_j(u) = u^T y_j$. Then, if $g_j(v) > 0$ let

(15) $h_j(w) = f_j(u)/g_j(v)$.

We say that w is D-improper if either

(16)   (a) $f_j(u) = 0$ for every DMU$_j$

or

(b) for some DMU$_j$, $f_j(u) > 0$, $g_j(u) = 0$.

Otherwise we call w D-proper.

If w is D-proper and if

(17) $h_j(w) \geq h_k(w)$ for all DMU$_k$ for which $g_k(v) > 0$

10

then we call w an optimizing multiplier for DMU_j and denote by W_j the set of all optimizing multipliers for DMU_j. We call W_j the multiplier set for DMU_j.

Notes: (a) If w is D-improper it is not assigned to any W_k. In Section 12 we will consider the significance of the D-improper virtual multipliers. (b) If $f_k(u) = g_k(v) = 0$ then $h_k(w)$ has the form 0/0 and is not only not defined but is not considered in or needed in applying (17).

(c) If $w = \begin{bmatrix} u \\ v \end{bmatrix} \in W_j$, then so is $w' = \begin{bmatrix} \alpha u \\ \beta v \end{bmatrix}$ for all $\alpha > 0$, $\beta > 0$.

W_j may be empty; if so we say that DMU_j is DEA-inefficient and denote by N (or N(D)) the set of all DEA-inefficient DMUs.

If W_j is not empty we say that DMU_j is DEA-scale-efficient and denote by RE (RE(D)) the set of all DEA-scale-efficient DMUs.

It follows readily from (17) that if D^* is a subdomain of D then

(18) $N(D^*) \subset N(D)$

and

(19) $N^* \cap RE(D) \subset RE(D^*)$.

Lemma 1. Suppose $k \neq h$, $P_k = \alpha P_h$ and let J^* be J with DMU_k deleted. Then

(20) $K(P^*) = K(P)$

(21) $N(D^*) = N(D) \cap N^*$

and

(22) $RE(D^*) = RE(D) \cap J^*$.

11

Proof: to establish (20) we observe that

$$P\lambda = P^*\lambda^*$$

where, for $j \neq h,k$, $\lambda_j^* = \lambda_j$ and $\lambda_h^* = \lambda_h + \alpha\lambda_k$. By the non-negativity of inputs and outputs, $\alpha > 0$ so that $\lambda_h^* \geq 0$. For (21) and (22) we note from (16)

(23) $h_k(w) = h_h(w)$ for all w, if $P_k = \alpha P_h$.

Corollary 2. Let D^* be the reduced subdomain obtained from D according to (14). Then (20), (21), and (22) hold for D^*.

The conclusions of Corollary 2 hold as well for any reduced subdomain D^* obtained from D by successive deletion of DMUs proportional to some remaining one.

It follows from Corollary 2 that for questions relating to scale efficiency or to the extended domain we may without loss of generality limit ourselves to reduced data domains.

Reduction assumption: Henceforth, unless otherwise specified, we assume that all data domains D considered are reduced.

4. CLASSIFICATION OF DEA EFFICIENCIES.

For our purposes, we make no distinction between two DMUs whose input-output vectors are equal. Hence we may identify a DMU with its input-output vector $p = \begin{bmatrix} y \\ -x \end{bmatrix}$.

Conceptually, any vector p with x and y non-negative and non-zero may be treated as a DMU. For any p in K(P) we construct a new data domain $D^+ = D^p$ with

(24) $J^+ = (P_1, \ldots, P_n, p)$ and $P^+ = [Pp]$.

(D^+ may not be reduced.)

We next define the _frontier_ FR of D as the set of all p in K(P) for which p is scale efficient in $D^+ = D^p$. Clearly,

(25) $RE \subset FR$.

In what follows we will give characterizations of FR in terms of RE.

We partition the set RE of scale efficient DMUs into three classes.

(26) $E = \{DMUk \in RE \mid \dim W_k = s + m\}$,

(27) $E' = \{DMUk \in RE \mid \dim W_k < s + m$ and

                there exists $w \in W_k$ such that $w > 0\}$,

(28) $F = \{DMUk \in RE \mid$ every $w \in W_k$ has at least one zero component$\}$.

We use the terms DEA-extreme-efficient, and DEA-non-extreme-efficient to describe, respectively, DMUs in E and E'. We do not consider elements of F to be DEA-efficient because (as will be seen below) the presence of zeros is related to the existence of slacks, which in turn indicate what has been called technical inefficiency. Thus, $E \cup E'$ is the set of all DEA-efficient DMUs, and F consists of the DMUs that are scale but not technically efficient.

5. SOME EXAMPLES.

Three related simple examples are displayed. Each colum. .ives the components y_{rj} and x_{ij} of the output and input vectors y_j and x_j of DMU_j. We recall that W is the set of all virtual multipliers

$$ w = \begin{bmatrix} u \\ v \end{bmatrix} , $$

where u is a non-negative, non-zero s vector and v is a non-negative, non-zero m vector.

In Example 1:

$$ W_1 = W, \quad W_2 = \{\text{all } w \text{ with } v_2 = 0\} $$

$\dim W_1 = 3 = m + s$ so that $DMU_1 \in E$

$\dim W_2 = 2 < m + s$ and has no positive vector so that $DMU_2 \in F$. E' and N are empty.

In Example 2:

$$ W_1 = \{\text{all } w \text{ with } v_3 \leq v_2\} $$

$$ W_2 = \{\text{all } w \text{ with } v_2 \leq v_3\} $$

Both W_1 and W_2 have dimension $4 = m + s$ and hence $E = \{DMU_1, DMU_2\}$. E', F, and N are empty. Note also that

$$ (29) \qquad W_1 \cap W_2 = \{\text{all } w \text{ with } v_2 = v_3\} $$

has dimension 3 which is less than $4 = m + s$.

In Example 3:

a and b are (positive) parameters and we distinguish three cases.

Case 3.1: $a + b < 1$. Then

$$ W_1 = \{\text{all } w \text{ with } v_3 \leq (a/(1-b))v_2\}, $$

$$ W_2 = \{\text{all } w \text{ with } v_3 \geq ((1-a)/b)v_2\}, $$

14

$$W_3 = \{\text{all } w \text{ with } (a/(1-b))v_2 \leq v_3 \leq ((1-a)/b)v_2\}.$$

Each of these has dimension $4 = m + s$ and hence $E = \{DMU_1, DMU_2, DMU_3\}$.
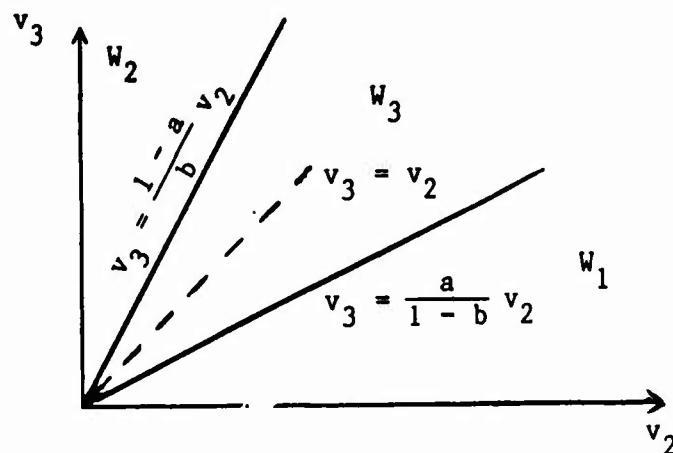
Figure 1: Projection on v_2, v_3 plane

Figure 1 is the projection of W onto the v_2, v_3 plane. Note that from $a + b < 1$ we conclude that $(1-a)/b > 1 > a/(1-b)$ and hence the multipliers w with $v_3 = v_2$ belong to W_3.

Case 3.2: $a + b = 1$.

The definitions for the W_i from case 3.1 are still valid but now $a/(1-b) = (1-a)/b = 1$. Hence, $\dim W_1 = \dim W_2 = 4$, $\dim W_3 = 3$ so that $E = \{DMU_1, DMU_2\}$, $E' = \{DMU_3\}$, and F and N are empty.

Note that as $a + b$ increases toward $a + b = 1$ in Case 3.1, the bounding lines in Figure 1 approach as a limiting position the line $v_2 = v_3$.

In Case 3.2

(30) $W_3 = W_1 \cap W_2$

and does not depend on the value of a (so long as $a + b = 1$). In this case we have

(31) $P_3 = bP_1 + aP_2.$

$$DMU_3(a) = (1-a)DMU_1 + aDMU_2$$
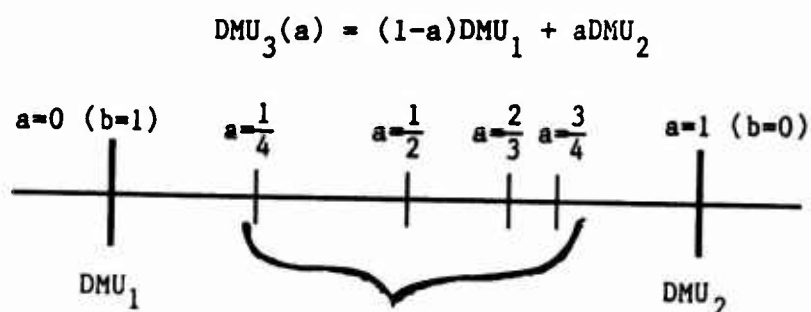
Figure 2: $DMU_3(a)$ for several values of a

Figure 2 illustrates P_3 for several values of a.

The independence of W_3 from a illustrates a property of the multipliers which requires special attention in some of the proofs which follow.

Case 3.3: $a + b > 1$.

In this case W_1 and W_2 are the same as in Case 3.2 but $W_3 = $ {all w with $v_2 = v_3 = 0$} has dim 2. Here $E = \{DMU_1, DMU_2\}$, $F = \{DMU_3\}$, and E' and N are empty. We can write

(32) $W_3 = W_1 \cap W_2 \cap$ {all w with $v_2 = v_3 = 0$}

and for $1-b \leq c \leq a$ we have

(33) $P_3 = (1-c)P_1 + cP_2 + R(c)$

where

(34) $R(c) = \begin{bmatrix} 0 \\ 0 \\ a-c \\ c+b-1 \end{bmatrix}$

16

The non-zero components of $R(c)$ are called "slacks" and measure the technical inefficiency of DMU_3. The lack of uniqueness of the expression (33) for P_3 and the consequential lack of uniqueness of the slack vector is significant in the theory of DEA-efficiency. [The fact that the sum $a + b - 1$ of the slacks in (34) is unique is not generally the case. For example, if we replace the value $x_{33} = 6$ with $x_{33} = 5.9$ the sum of the slacks $a + b - 1 + (0.1)c$ is not independent of c.]

Table 4 gives an example which shows that DEA-efficiency, by itself, may not always guarantee desirability. In this example with $m = 4$, $s = 3$, $n = ms = 12$ (clearly generalizable for any m and s) every DMU_j is extremely efficient. To see this write j in the form

$$j = (\alpha - 1)s + \beta$$

where $1 \leq \alpha \leq m$ and $1 \leq \beta \leq s$. Then, if we define w_j by

$$(35) \qquad u_{rj} = \delta_{r\beta} \qquad (r = 1,\ldots,s),$$

$$v_{ij} = \delta_{i\alpha} \qquad (i = 1,\ldots,m)$$

where δ_{pq} is the kronecker delta, we have

$$(36) \qquad h_j(w_j) = 3 > h_k(w_j) \text{ for all } k \neq j.$$

From (36) it follows at once that $DMU_j \in RE_j$; however, from Theorem 6(E2) proved later, the "greater than" in (36) implies also that $DMU_j \in E$, i.e., that dim $W_j = s + m = 7$.

	D₁	
	DMU	
	1	2
y_{1j}	1	1
x_{1j}	1	1
x_{2j}	5	6

Example 1: $m = 2$, $s = 1$, $n = 2$

	D₂	
	DMU	
	1	2
y_{1j}	1	1
x_{1j}	1	1
x_{2j}	5	6
x_{3j}	6	5

Example 2: $m = 3$, $s = 1$, $n = 2$

	DMU		
	1	2	3
y_{1j}	1	1	1
x_{1j}	1	1	1
x_{2j}	5	6	$5 + a$
x_{3j}	6	5	$5 + b$

Example 3: $m = 3$, $s = 1$, $n = 3$, $0 < a \leq 1$, $0 < b \leq 1$

| | DMU | | | | | | | | | | | |
	1	2	3	4	5	6	7	8	9	10	11	12
output 1	3	2	2	3	2	2	3	2	2	3	2	2
2	2	3	2	2	3	2	2	3	2	2	3	2
3	2	2	3	2	2	3	2	2	3	2	2	3
input 1	1	1	1	2	2	2	2	2	2	2	2	2
2	2	2	2	1	1	1	2	2	2	2	2	2
3	2	2	2	2	2	2	1	1	1	2	2	2
4	2	2	2	2	2	2	2	2	2	1	1	1

Example 4: an example with every DMU
extremely efficient

5. A LINEAR PROGRAMMING FORMULATION.

Following CCT [16] we next turn to a linear programming formulation of DEA-efficiency. Let α and β be positive scalars and let

(37) $\qquad w' = \begin{bmatrix} \alpha u \\ \beta v \end{bmatrix}$

Lemma 3. For any k, $w \in W_k$ iff $w' \in W_k$.

This lemma follows from (16) and (17).

In view of Lemma 3 the determination as to whether or not DMU_j is DEA-scale-efficient is equivalent to the existence or non-existence of a virtual multiplier w for which

(38) $\qquad 1 = h_j(w) \geq h_k(w)$ for all DMU_k for which $g_k(v) > 0$

or even, indeed, for which in addition to (38)

(39) $\qquad g_j(v) = 1.$

Of course (38) and (39) also imply

(40) $\qquad f_j(u) = 1$

For any j we denote by W_j^m the set of all w (if any) which satisfy (38) and (39). We call W_j^m the normalized multiplier set for DMU_j. Clearly, if w' and w satisfy (37) then

19

(41) $W_j = \{w \mid \text{there exist } \alpha \text{ and } \beta \text{ for which } w' \in W_j^m\}$.

Let "o" stand for some integer from $\{1,\ldots,m\}$. We associate with DMU_o the scalar z^* defined by the non-linear optimization problem:

(42) $z^* = \max \{h_o(w) \mid D\text{-proper } w \text{ for which } h_j(w) \leq 1 \text{ for all } j \text{ with}$

$$g_j(v) > 0\}.$$

Clearly $z^* = 1$ iff DMU_o is scale efficient. Moreover, if $z^* = 1$ there is a $w \in W_o^m$ which solves (42), i.e., which satisfies (39) for DMU_o.

Since, when $g_j(v) > 0$

(43) $h_j(w) \leq 1$ iff $f_j(u) - g_j(v) \leq 0$

and since $f_j(u) - g_j(v) = 0 - 0 \leq 0$ for any D-proper w with $g_j(v) = 0$ we can replace the non-linear system (42) and (39) by the linear program

(44) $\max \quad z = \sum_r y_{ro} u_r$

subject to

$$\sum_r y_{rj} u_r - \sum_i x_{ij} v_i \leq 0, \quad j=1,\ldots,n$$

$$\sum_i x_{io} v_i = 1$$

$$u_r \geq 0 \qquad r=1,\ldots,s$$

$$v_i \geq 0 \qquad i=1,\ldots,m$$

This program may be regarded as the dual of a related linear program which we will reference as the primal problem. The primal and dual problems are shown in Table 1.

20

primal	dual

<div style="display:flex">

primal

$\min \quad \theta$

s.t.

$\quad \theta \ (\text{free})$

$\quad \lambda_j \geq 0$

$x_{io}\theta - \sum_j x_{ij}\lambda_j \geq 0$

$\sum_j y_{rj}\lambda_j \geq y_{ro}$

dual

$\max z = \sum_r y_{ro}u_r$

s.t.

$\sum_i x_{io}v_i = 1$

$\sum_r y_{rj}u_r - \sum_i x_{ij}v_i \leq 0 \quad j=1,\ldots,n$

$u_r \geq 0 \qquad r=1,\ldots,s$

$v_i \geq 0 \qquad i=1,\ldots,m$

</div>

Table 1 - Primal and Dual Programs for DEA Scale Efficiency

In Table 1 the vector $\lambda^{\#}$ defined by $\lambda_o^{\#} = 1$, $\lambda_j^{\#} = 0$ for $j \neq 0$, and corresponding $\theta^{\#} = 1$ is always feasible. On the other hand, by hypothesis, there is at least one i for which $x_{io} > 0$. Then the primal relation

$$x_{io}\theta - \sum_j x_{ij}\lambda_j \geq 0$$

requires the θ be non-negative for any feasible (λ,θ). Hence, the primal problem will have an optimal feasible solution (λ^*,θ^*) with $0 \leq \theta^* \leq 1$.

We can strengthen this result to

$$0 < \theta^* \leq 1$$

For, let (λ,θ) be feasible for the primal. Then $\lambda \neq 0$ else

$$0 = \sum_j y_{rj}\lambda_j \geq y_{ro} \text{ for all } r$$

contrary to the hypothesis that y_{ro} is positive for at least one r. Suppose, say, that $\lambda_k > 0$, then for every i

$$\theta x_{io} \geq \sum_j x_{ij}\lambda_j \geq x_{ik}\lambda_k.$$

Since, by hypothesis for at least one i we have $x_{ik} > 0$ we conclude $\theta x_{io} > 0$ and hence $\theta > 0$. In particular, $\theta^* > 0$.

21

By duality, the dual problem will also have one or more optimal solutions all with $0 < z^*(=\theta^*) \leq 1$. This fact and the condition $g_0(v) \doteq 1$ together show that <u>every optimal solution w^* of the dual is D- proper.</u>

We associate a slack with each inequality in the primal or the dual:

$$(45) \qquad s_i^- = x_{io}\theta - \sum_j x_{ij}\lambda_j \geq 0 \qquad (i = 1,\ldots,m)$$

$$(46) \qquad s_r^+ = -y_{ro} + \sum_j y_{rj}\lambda_j \geq 0 \qquad (r = 1,\ldots,s)$$

$$(47) \qquad t_j = -\sum_r y_{rj}u_r + \sum_i x_{ij}v_i \geq 0 \qquad (j = 1,\ldots,n)$$

By the complementary slackness theorem for dual programs we have for any optimal solutions λ and w of the primal and dual, respectively,

$$(48) \qquad s_i^- v_i = s_r^+ u_r = t_j\lambda_j = 0 \qquad (i = 1,\ldots,m; \; r = 1,\ldots,s;$$
$$j = 1,\ldots,n).$$

If the optimal values $\theta^* = z^*$ are less than 1 then by (47) the dual slack t_o is positive so that $\lambda_o = 0$ for all optimal λ. However, if $\theta^* = z^* = 1$ then $\lambda^\#$ is one optimal solution for the primal, and for it the primal slacks are all zero.

A less well known feature of duality known as the <u>strong complementary slackness theorem</u>[1] states that there exists an optimal solution pair (λ^o, w^o) for which, in addition to (48), we have

$$(49) \qquad s_i^{o-} + v_i^o > 0 \qquad (i = 1,\ldots,m)$$
$$s_r^{o+} + u_r^o > 0 \qquad (r = 1,\ldots,s)$$
$$t_j^o + \lambda_j^o > 0 \qquad (j = 1,\ldots,n)$$

Next, suppose that DMU_o is scale efficient, i.e., for the optimal solutions (λ^o, w^o) we have $\theta^o = z^o = 1$. Then, from (45) and (46) we get, in vector form,

$$(50) \qquad x_o = \sum_j \lambda_j^o x_j + s^{o-}$$

[1]See Section 13 below for discussion and references.

22

$$(51) \qquad y_o = \sum_j \lambda_j^o \, y_j - s^{o+}$$

$$(52) \qquad P_o = \sum_j \lambda_j^o \, P_j - s^{\neq o} = \sum_j \lambda_j^o \begin{bmatrix} y_j \\ -x_j \end{bmatrix} - \begin{bmatrix} s^{o+} \\ s^{o-} \end{bmatrix} = P_o^o - s^{\neq o}$$

where $s^{\neq} = \begin{matrix} s^+ \\ s^- \end{matrix}$ is the non-negative vector of all primal slacks, and P_o^o defines a DMU in the conical hull $K(P)$ of D.

If s^{+o} is non-zero then (52) states that there is a DMU in $K(P)$ with the same or less input and more output. Similarly, if s^{-o} is non-zero, then (52) states that there is a DMU in $K(P)$ with the same or more output and less input.

In either case it is not logical to regard DMU_o as being fully efficient. Hence, if $s^{\neq o}$ is non-zero we do not call DMU_o DEA-efficient even though it is DEA-scale-efficient. We use the term technical efficiency to indicate $s^{\neq o} = 0$, and use the term DEA-efficient to describe a DMU for which $z^o = 1$ and $s^{\neq o} = 0$, i.e., one that is both scale efficient and technically efficient.

Note that we do not preclude the consideration of stronger forms of efficiency than DEA-efficiency. Thus, whereas to be DEA-inefficient is a mark against any DMU, to be DEA-efficient need not be enough to settle efficiency in some contexts. We do not treat the question of general efficiency in this paper.

Suppose that DMU_o is DEA-efficient. Then from (49) we conclude that every component of w^o is positive. Conversely, if W_o contains any positive vector w then from (48) applied to λ^o and w we conclude that $s^{\neq o} = 0$ and hence DMU_o is DEA-efficient.

23

7. THE PRINCIPAL CLASSIFICATION THEOREMS.

We begin with a few more definitions. For any j we define a set $D(j)$ called the dominant of DMU_j by

$$(53) \qquad D(j) = \{DMU_k | W_j \subset W_k\}$$

and then also introduce

$$(54) \qquad E(j) = D(j) \cap E$$

Note, in particular, that if $DMU_j \in N$ then $D(j) = J$ and $E(j) = E$.

Next, for any virtual multiplier vector $w \in W$ we define another subset of J by

$$(55) \qquad J(w) = \{DMU_j | w \in W_j\}.$$

Referring again to the dual optimal vectors λ^o, w^o, of the previous section we define partitions J^+, J^o; I^+, I^o; R^+, R^o of J, $I = \{1,\ldots,m\}$, $R = \{1,\ldots,s\}$, respectively, by

$$(56) \qquad J^+ = \{DMU_j | \lambda_j^o > 0\}, \quad J^o = \{DMU_j | \lambda_j^o = 0\}$$
$$I^+ = \{i \quad | v_i^o > 0\}, \quad I^o = \{i \quad | v_i^o = 0\}$$
$$R^+ = \{r \quad | u_r^o > 0\}, \quad \bar{R}^o = \{r \quad | u_r^o = 0\}.$$

Then from (51) through (55) we get

$$(57) \qquad t_j^o = -\sum_r y_{rj} u_r^o + \sum_i x_{ij} v_i^o \quad \begin{cases} > 0 \text{ for } DMU_j \in J^o \\ = 0 \text{ for } DMU_j \in J^+ \end{cases}$$

$$(58) \qquad s_i^o = x_{io} o^o - \sum_j x_{ij} \lambda_j^o \quad \begin{cases} > 0 \text{ for } i \in I^o \\ = 0 \text{ for } i \in I^+ \end{cases}$$

$$(59) \qquad s_r^{o+} = -y_{ro} + \sum_j y_{rj} \lambda_j^o \quad \begin{cases} > 0 \text{ for } r \in R^o \\ = 0 \text{ for } r \in R^+ \end{cases}$$

For any subsets J', I', R', of J, I, R, respectively, we set

$$(60) \qquad \lambda(J') = \{\lambda | \lambda_j = 0 \text{ if } DMU_j \in J'\}$$

$$(61) \qquad W(R',I') = \{w = (u,v) \in W \begin{vmatrix} u_r = 0 & r \in R' \\ v_i = 0 & i \in I' \end{vmatrix}\}$$

$$(62) \qquad P(R',I') = \{p = \begin{bmatrix} y \\ -x \end{bmatrix} \begin{vmatrix} y_r = 0 & r \in R' \\ x_i = 0 & i \in J' \end{vmatrix}\}$$

24

We will say that a vector w is <u>maximally positive in</u> $W(R',I')$ if $w \in W(R',I')$ and

(63) $u_r > 0$, $v_i > 0$ for all $r \notin R'$, $i \notin I'$,

i.e., all of the components of w which are not required to be zero are actually positive. Note that in view of (49) the slack vector $s^{\bar{*}0}$ in (52) is maximally positive in $W(R^+,I^+)$.

Theorem 4. Primal Representation Theorem.

Suppose that $DMU_o \in RE$. Then we can write

(64) $\qquad P_o = \sum_{DMU_j \in E(o)} \alpha_j P_j - s^{\ddagger o}$

where $\alpha_j \geq 0$ for all j and $s^{\ddagger o}$ is a slack vector which is maximally positive in $W(R^+,I^+)$.

Neither the coefficients α_j nor the slack vector $s^{\ddagger o}$ are necessarily unique.

Theorem 5. Dual Representation Theorem.

Suppose that $DMU_o \in RE$. Then

(65) $\qquad W_o = \bigcap_{DMU_j \in E(o)} W_j \cap W(R^o,I^o)$

Theorem 6. DMU_o belongs to E if and only if any one of the following properties holds:

(E1) $\qquad \dim W_o = s + m$,

(E2) $\qquad$ for some $w \in W_o$, $J(w) = \{DMU_o\}$,

(E3) $\qquad \dim W_o \cap W_j = s + m$ iff $j = 0$,

(E4) $\qquad$ the primal program (cf Table 1) has the unique optimal solution

$\qquad\qquad \lambda_o = 1$, $\lambda_j = 0$ for all $j \neq 0$.

(E5) $\qquad D(o) = \{DMU_o\}$

(E6) $\qquad J^+(o) = \{DMU_o\}$

Theorem 7. DMU_o belongs to E' if and only if any one of the following properties holds:

(E'1)   dim $W_o < s + m$ and $w^o > 0$.

(E'2)   the primal program (cf Table 1) has more than one optimal solution
          and $w^o > 0$.

(E'3) W_o is non empty, $s^{\dagger o} = 0$ (in(52)), and dim $W_o < s + m$

(E'4) W_o is non empty, $s^{\dagger o} = 0$ (in Theorem 4), and dim $W_o < s + m$

(E'5) W_o is non empty, contains at least one positive vector, and
          dim $W_o < s + m$

Theorem 8. DMU_o belongs to F if and only if any one of the following
conditions holds:

(F1) W_o is non empty but contains no positive vector.

(F2) W_o is non empty and w^o has at least one zero conponent.

(F3) W_o is non empty and $s^{\dagger}(o) \neq 0$ (in Theorem 4)

(F4) W_o is non empty and $s^{\dagger o} \neq 0$ (in 52).

We will prove these theorems in the next three sections.

8. PROOFS OF THE PRINCIPAL CLASSIFICATION THEOREMS (PART 1).

In this section we proceed under the hypothesis that DMU_o is scale
efficient. Hence from (52) and (56) we have

(66) $P_o = \sum_{DMU_j \epsilon J^+} \lambda_j^o P_j - s^{\dagger o}$

          where $\lambda_j^o > 0$ and $s^{\dagger o}$ is maximally positive in $W(R^+, I^+)$

          <u>Lemma 9</u>. $J^+ = D(o) = J(w_o)$.

Proof: suppose that $\lambda_k^o = 0$, then $t_k^o > 0$ and from (57) we conclude that
$\lambda^o \notin W_k$ and hence $DMU_k \notin D(o)$. Hence $J^+ \supset D(o)$.

On the other hand, if $DMU_k \epsilon J^+$ then $\lambda_k^o > 0$. Now for $w \epsilon W_o^m$ we get
from (66) that $0 = w^T P_o = \sum_{DMU_j \epsilon J^+} \lambda_j^o w^T P_j - w^T s^{\dagger o}$.

From (60) with $\lambda = \lambda^o$ we get $w^T s^{\dagger o} = 0$. Next, since in each product $\lambda_j^o w^T P_j$ the first factor is positive and the second is non-negative, for the sum to be zero each term $w^T P_j = 0$ and therefore $D(o) \supset J^+$.

Finally, $DMU_j \in J(w^o)$ if and only if $t_j = 0$ and hence, by (49), if and only if $\lambda_j^o > 0$. Thus (cf (56)) $J^+ = J(w^o)$.

The proof of the next lemma utilizes a well-known property of analytic functions which when applied to two quotients h_o and h_j of homogeneous linear functions, of $(s + m)$ variables states that equality of h_o and h_j over an $(s + m)$ sphere implies that there is a positive scalar α for which the numerator and denominator of h_j are respectively just α times the numerator and denominator of h_o.

Lemma 10. Let $DMU_o \in E$. Then $\dim W_o \cap W_j = s + m$ iff $j = o$.
Proof: The "if" follows immediately from the definition of E. For the "only if" we have $h_o(w) = u^T y_o / v^T x_o$, $h_j(w) = u^t y_j / v^T x_j$. Now, if $\dim W_o \cap W_j = s + m$ then $W_o \cap W_j$ contains an $(s + m)$ sphere on which $h_o = h_j$. It follows from the cited property of analytic functions that for some positive scalar α we have $y_j = \alpha y_o$ and $x_j = \alpha x_o$ and hence $P_j = \alpha P_o$. Since our data domain D is reduced this requires $\alpha = 1$ and $j = o$.

Lemma 11. $D(o) = \{DMU_o\}$ iff $DMU_o \in E$
Proof: If $\{DMU_o\} = D(o) = J(w^o)$ then W_o contains all elements of W sufficiently close to w^o and hence has dimension $s + m$ so that $DMU_o \in E$. Conversely, if $DMU_o \in E$ then for any $w \in W_o^n$ we have from (52) and the definition of W_o^n that

(67) $0 = w^T P_o = \sum_j \lambda_j^o (w^T P_j) - w^T s^{\dagger o}$.

From (48) we have $w^T s^{\dagger o} = 0$ (whether or not $DMU_o \in E$). From (44) we have $w^T P_j \leq 0$ and hence $\lambda_j^o > 0$ only if $w^T P_j = 0$. Thus if $\lambda_j^o > 0$ we have

27

$W_o \subset W_j$ and hence dim $W_o \cap W_j$ = dim W_o = s + m. Now by Lemma 10, we must have j = o.

Lemma 12. If DMU_o does not belong to E, then $\lambda_o^o < 1$ and, hence,

(68) $P_o = \sum\limits_{DMU_j \, \epsilon \, J^+ - \{DMU_o\}} \rho_j P_j - s^{\dagger'}$

where $\rho_j = \lambda_j/(1-\lambda_o)^{-1}$ and where

(69) $s^{\dagger'} = (1-\lambda_o)^{-1} s^{\dagger o}$

is maximally positive in $W(R^+, I^+)$

Proof: from (66) we have

(70) $(1-\lambda_o^o)x_o = \sum\limits_{DMU_j \, \epsilon \, J^+ - \{DMU_o\}} \lambda_j^o x_j - s^{-o}.$

Since $DMU_o \notin E$, $J^+ - \{DMU_o\}$ is non-empty (by Lemma 11). Now, since each λ_j^o in the summation is positive, since each x_j has no negative and at least one positive component, and since $R^{o-} \geq 0$ the right hand side of (70) is a non-negative non-zero vector. Since $x_o \geq 0$ this requires $1 - \lambda_o^o > 0$.

Lemma 13. Let λ^* be defined by $\lambda_o^* = 1$, $\lambda_j^* = 0$ for all j ≠ o. If $DMU_o \, \epsilon \, E$ then $\lambda^* = \lambda^o$ and is the unique optimal solution of the primal problem. Moreover, $s^{\dagger o} = 0$ and $R^+ = R$, $I^+ = I$.

In the proof of Lemma 11 we established that $\lambda_j^o = 0$ for all j ≠ o. Thus (48) reduces to

(71) $P_o = \lambda_o^o P_o - s^{\dagger o}$

so that $\lambda_o^o = 1$ and $s^{\dagger o} = 0$. Next, from (58) and (59) we conclude that $R^+ = R$ and $I^+ = I$ (and, of course, that both R^o and I^o are empty).

At this stage we have established Theorems 4 and 5 for $DMU_o \, \epsilon \, E$. For, E(o) = D(o) = $\{DMU_o\}$ and $W(R^o, I^o) = W$, (71) reduces to (64), and (65) reduces to $W_o = W_o \cap W$.

Lemma 14. If the primal system has a unique optimal solution then that solution is the λ^* described in Lemma 13 and $DMU_o \, \epsilon \, E$.

28

Proof: Since λ^* is an optimal solution whenever $DMU_o \in RE$ the unique optimal solution must be λ^*. Then $\lambda^o = \lambda^*$ and $D(o) = J^+ = \{DMU_o\}$. Now by Lemma 11 we conclude that $DMU_o \in E$.

The general framework of the proof of Theorem 4 is to begin with (68) and set up a step-by-step reduction of the summation range from $J^+ - DMU_o$ to $E(o)$. This would be trivial if W_j properly contained W_o for every DMU_j in $J^+ - \{DMU_o\}$. The following Example 3.4 (which is derived from the earlier Example 3) shows that we can have $W_j = W_o$ for $j \neq 0$ and $DMU_j \in J^+$ and thus illustrates the need for the constructions introduced below when we pick up the proof of Theorem 4.

	DMU
	1	2	3	4	5	6
y_{1j}	1	1	1	1	1	1
x_{1j}	1	1	1	1	1	1
x_{2j}	5	6	5.2	5.4	5.6	5.8
x_{3j}	6	5	5.8	5.6	5.4	5.2

Example 3.4. Expansion of
Example 3. $s = 1$, $m = 3$, $n = 6$.

We have

$$W_1 = \{\text{all } w \text{ with } v_3 \leq v_2\}$$

$$W_2 = \{\text{all } w \text{ with } v_2 \leq v_3\}$$

$$W_3 = W_4 = W_5 = W_6 = \{\text{all } w \text{ with } v_2 = v_3\} = W_1 \quad W_2$$

$$\dim W_1 = \dim W_2 = 4 = s + m$$

$$\dim W_3 = \dim W_4 \ \dim W_5 = \dim W_6 = 3 < s + m$$

Therefore, $E = \{DMU_1, DMU_2\}$, $E' = \{DMU_3, DMU_4, DMU_5, DMU_6\}$. Now let $o = 4$; we

29

can verify that $\lambda^o = \frac{1}{60}(13,7,5,5,5,5)$ and $w^o = \begin{bmatrix} u^o \\ v^o \end{bmatrix} = \frac{1}{12}\begin{bmatrix} 12 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ are op-

timal primal and dual solutions that satisfy the strong complementary

slackness property. Next, the equation

(64') $P_4 = (0.6)P_1 + (0.4)P_2 - 0$

meets the conditions of Theorem 4 since $E(o) = \{DMU_1, DMU_2\} = E$ and $R(o) = 0$

is maximally positive in $W(R^+, I^+)$ since $R^+ = R$ and $I^+ = I$.

On the other hand, $J^+ = J$ and equation (66) becomes

(66') $P_o = \frac{1}{60}(13P_1 + 7P_2 + 5P_3 + 5P_4 + 5P_5 + 5P_6) - 0$

so that (67) becomes

(67') $P_o = \frac{1}{55}(13P_1 + 7P_2 + 5P_3 + 5P_4 + 5P_5 + 5P_6) = 0.$

What then is our logical path from (66') to (64')? One possible approach is

to show that P_o can be written in a form such as (67') but where for each P_j

on the right hand side W_j properly contains W_{oj} i.e., in the present case in

terms of P_1 and P_2. We turn to this task in the following section.

9. PROOFS OF THE PRINCIPAL CLASSIFICATION THEOREMS (PART 2).

We begin by recalling the expression

(68) $P_o = \sum_{DMU_j \, \epsilon \, J^+ - DMU_o} \rho_j P_j = -s^{+'}$

where $s^{+'}$ is maximally positive in $W(R^+, I^+)$ (Recall that $DMU_o \, \epsilon \, RE - E.)$

Let $T = \{DMU_j | j \, \epsilon \, J - \{DMU_o\}$ and $W_j = W_o\}$.

Lemma 15. Let $J^* = J - T$, let P^* be obtained from P by deleting all

columns P_j for which $DMU_j \, \epsilon \, T$, and let $o^* = o$. Then $W_o^* = W_o$.

Proof: If T is empty there is nothing to prove. Suppose that T is non-

empty (it follows from Lemma 10 that this cannot happen if $DMU_o \, \epsilon \, E$). Since

every condition for membership in W_o^* holds also for W_o (cf (17)), $W_o \subset W_o^*$.

30

Suppose then that $w' \in W_o^*$ and $w' \notin W_o$. By the definition of w' we have

(72) $h_j(w') \leq h_o(w')$ for all $DMU_j \in J^*$.

Since $w' \notin W_o$ there is some $DMU_k \in T$ for which $h_o(w') < h_k(w')$, and if there are several such DMU_k we choose one for which

(73) $h_j(w') \leq h_k(w')$ for all $DMU_j \in T$.

Now, since $J = J^* \cup T$, we conclude from (72) and (73) that $w' \in W_k$ contrary to our hypothesis that

$W_k = W_o$.

Lemma 16. For any $DMU_o \in RE - E$ there is a expression of the form

(74) $P_o = \sum\limits_{j \in J^+-T-\{DMU_o\}} \rho_j P_j - s^{+''}$

where $\rho_j \geq 0$ and $s^{+''}$ is maximally positive in $W(R^+,I^+)$, and for each P_j appearing on the right hand side W_j properly contains P_o.

Proof: By the definition of T for each $DMU_j \in J^+ - T - \{DMU_o\}$ we do have W_j properly containing W_o and hence also W_j^* properly contains $W_o^* = W_o$. Next applying our generaly theory to J^* we get an expression of the form (74) but cannot assume that the slack vector $s^{+''}$ is maximally positive in $W(R^+,I^+)$.

However, since $w^o \in W_o^*$ we do have that $s^{+''} \in W(R^+,I^+)$. Moreover such an expression holds for each $DMU_k \in T$, i.e.,

(75) $P_k = \sum\limits_{DMU_j \in J^+-T-\{DMU_o\}} \rho_{jk} P_j - s_k^{+''}$, $DMU_k \in T$.

Next substitute (75) in (67) for each $k \quad T$ and we get

(76) $P_o = \sum\limits_{DMU_j \in J^+-T-\{DMU_o\}} \rho_j' P_j - s^{+''}$

where

$\rho_j' = \rho_j + \sum\limits_{DMU_k \in T} \rho_j k \geq 0$ for all $DMU_j \in J^+ - T - \{DMU_o\}$

and where

$$s^{\dagger''} = s^{\dagger'} + \sum_{DMU_k \, \varepsilon \, T} s_k^{\dagger''}$$

belongs to $W(R^+,I^+)$ since each summand does and is maximally positive in $W(R^+,I^+)$ since $s^{\dagger'}$ is.

Lemma 17. For any scale efficient DMU_o we have

$$N^+ = \{DMU_k | W_k \supset W_o\}$$

and for and $DMU_k \subset J^+$ we have

$$J(k)^+ \subset J^+, \; R(k)^+ \supset R^+, \; I(k)^+ \supset I^+$$

and

$$W(R(k)^+, I(k)^+) \subset W(R^+, I^+).$$

Proof: If $DMU_k \, \varepsilon \, J^o$ then we have $\sum y_{rk}u_r^o - \sum x_{ik}v_i^o < 0$
so that W_o is not a subset of W_k. On the other hand, if $w \, \varepsilon \, W_o^n$ and
$DMU_k \, \varepsilon \, J^+$, then from $\lambda_k^o > 0$ we conclude (via complementary slackness of λ^o and w) that

$$\sum y_{rk}u_r - \sum x_{ik}v_i = 0;$$

hence $w \, \varepsilon \, W_k$ and consequently $W_o \subseteq W_k$.

Next, suppose that $W_k \supset W_o$. Then, if $W_j \supset W_k$, we conclude that
$W_j \supset W_o$; hence $J(k)^+ \subset J^+$. Moreover, if $r \, \varepsilon \, R(k)^o$, $i \, \varepsilon \, I(k)^o$ then for any
$w \, \varepsilon \, W_k^n$ we have $u_r = v_i = 0$. In particular, $w^o \, \varepsilon \, W_k^n$ so that $u_r^o = v_i^o = 0$ and
hence $R(k)^o \supset R^o$, $I(k)^o \supset I^o$. It follows from (56) that $R(k)^+ \subset R^+$,
$I(k)^+ \subset I^+$ and hence also that $W(R(k)^+, I(k)^+) \subset W(R^+, I^+)$.

We are now ready to complete the proof of Theorem 4.

Suppose that $DMU_o \, \varepsilon \, RE - E$ and we have an expression for P_o of the form

(77) $$P_o = \sum_{DMU_j \, \varepsilon \, K} \lambda_{Kj}P_j - s_K^{\dagger} \quad \text{where}$$

(78a) $E(o) \subset K \subset J^+ - T - \{DMU_o\}$

(78b) $\lambda_{kj} > 0$ for all $DMU_j \, \varepsilon \, K$ and

(78c) s_K is maximally positive in $W(R^+, I^+)$.

Next, suppose $W_k \supset W_j$ for all $DMU_j \in K$ and let

$$T(k) = \{DMU_j | W_j = W_k \text{ and } DMU_j \in K - \{DMU_k\}\}.$$

Now apply Lemma 16 for all $o = j$ where

$DMU_j \subset T(k) \cup \{DMU_k\}$, substitute the resulting expressions on the right side of (77) and we will get an expression of the form (77) where the summation is over a set K' such that (78a,b,c) all hold and for every $DMU_j \in K'$, W_j properly contains W_k. This process can be iterated until we reach a K^p which is equal to $E(o)$. This completes the proof of Theorem 4.

Proof of Theorem 5: Since $E(o) = D(o) \cap E$ and $D(o) = \{DMU_k | W_o \subset W_k\}$ we know that $W_o \subset {}_{DMU_j \in \widehat{E}(o)} W_j$. Moreover, since for each $w \in W_o^n$ we have from (56) with $\lambda = \lambda^o$ that $w \in W(R^o, I^o)$. Therefore

$$W_o \subset {}_{DMU_j \in \widehat{E}(o)} W_j \cap W(R^o, I^o).$$

Next, suppose that P_o satisfies (64) and that w belongs to the right hand side of (65). By the definition of $E(o)$, all $h_j(w)$ for $DMU_j \in E(o)$ are equal. Let the common value be β. Then, $w' = \begin{bmatrix} u \\ \beta v \end{bmatrix}$ also belongs to the right hand side of (65) and for it

$$h_j(w') = 1 \geq h_k(w') \text{ for all } DMU_j \in E(o), DMU_k \in J. \text{ Multiply (64)}$$

by w'^T and we get

$$u^T y_o - \beta v^T x_o = w'^T P_o = \sum_{DMU_j \in E(0)} \alpha_j w'^T P_j - w'^T s^{+o}$$

$$= \sum \alpha_j 0 - 0 = 0.$$

Therefore, $h_o(w') = 1$ and so w' and hence also w is in W_o. This shows that

$$W_o \supset {}_{DMU_j \in E(0)} W_j \cap W(R^o, I^o)$$

and completes the proof of Theorem 5.

33

10. PROOFS OF THE PRINCIPAL CLASSIFICATION THEOREMS (PART 3)

Proof of Theorem 6: The lemmas in Section 8 take care of most of theorem 6.

(E1) is the definition of E

(E2) follows from Lemmas 9 and 11

(E3) is Lemma 10

(E4) implies (E6) and therefore (E1). Lemma 13 implies (E4)

(E5) is Lemma 11

(E6) is equivalent to (E5) by Lemma 9.

Proof of Theorem 7:

(E'1) is the definition of E'

(E'2) follows from Theorem 5

(E'3) If W_o is non-empty and $s^{\dagger o} = 0$, then since $s^{\dagger o}$ is maximally positive in $W(R^+, I^+)$ we must have $R^+ = R$, $I^+ = I$ and hence $w^o > 0$. Therefore, (E'3) implies (E'1). The reverse implication is obvious from (49).

(E'4) A similar argument to the one for (E'3) holds for (E'4).

(E'5) If W_o contains any positive vector w then w^o is positive and $s^{\dagger o} = 0$ so that (E'5) implies (E'3). Conversely, if $s^{\dagger o} = 0$ then w^o is positive so that (E'3) implies (E'5).

Proof of Theorem 8:

(F1) is the definition of F.

(F2) If w^o has a zero component then $s^{\dagger o}$ has a positive component and therefore no $w \in W_o$ can be positive. Hence (F2) implies (F1). The converse is trivial.

(F3) and (F4) both imply that W_o can contain no positive vector and hence that (F1) holds. Next, if w^o is not positive then $s^{\dagger o} \neq 0$ so that (F1)

34

implies (F4). Finally, if w^o is not positive then we cannot have both $R^+ = R$ and $I^+ = I$ so that $s^\dagger(o) \neq 0$. Therefore (F1) implies (F3).

11. CLASSES OF NON SCALE EFFICIENT DMUs

Suppose that $DMU_o \in N$ and that λ^*, w^* are optimal dual vectors for DMU_o. Then consider a new DMU_o' defined by

$$P_o' = \begin{bmatrix} y_o \\ -\theta^* X_o \end{bmatrix} \ .$$

(It may indeed happen that for some j and $\alpha > 0$ we have $P_o' = \alpha P_j$. This makes no essential difference in the discussion which follows.)

From the primal equations we have

$$(79) \qquad P_o' = \sum_{j=1}^n \lambda_j^* P_j - s^{\dagger *} \leq \sum_{j=1}^n \lambda_j^* P_j .$$

Since $DMU_o \in N$, $t_o^* > 0$ and hence $\lambda_o^* = 0$, therefore (79) shows that $P_o' \in K(D)$ (the conical null of D). Moreover, (79) shows that DMU_o' is scale efficient in the domain D^+ (for $p = P_o'$) and hence, that DMU_o' is in the frontier FR of D, and therefore belongs to (exactly) one of the classes E, E', F in D^+. Based on this we partition N into three classes as follows:

$$(80) \qquad DMU_o \in NE \text{ or } NE' \text{ or } NF \text{ according as } DMU_o \in E \text{ or } E' \text{ or } F \text{ in } D^+.$$

We now have a partition of the DMUs into six classes.

For randomly generated data we might conjecture that the classes E', F and IE are very small as contrasted to E, IE' and I. It would be interesting to check this conjecture for data domains which have been studied. (The 15 hospital case in [6] was contrived to enhance the likelihood that F would not be empty, but what about e' and IE?)

35

12. THE ROLE OF ZERO INPUTS AND OUTPUTS

We have established the classification theorems for DMUs without the earlier DEA assumptions that all x_{ij} and y_{rj} were non-zero. However, further discussion of the influence of zero inputs and outputs is warranted. In particular, it follows from the universal feasibility of the programs in Table 1 that no DMU_j can become scale efficient only because of some w with $\sum x_{ij} v_j = 0$.

In discussing zero inputs it is convenient to introduce subsets $I_1, \ldots, I_n$ of I by the definition

(81) $\qquad I_j = \{i \mid x_{ij} = 0\}, \; j = 1, \ldots, n.$

Theorem 18. Let I' be a non-empty subset of I for which

(82) $\qquad I' \subset I_j \;\; \text{iff} \;\; j \; \varepsilon \; \{1, \ldots, k\}$

and let D^* be the reduced data domain consisting of $DMU_1, \ldots, DMU_k$. Then if $o \; \varepsilon \; \{1, \ldots, k\}$ the category of DMU_o is the same in D^* as in D.

Consider the primal equation

(83) $\qquad \partial x_{io} = \sum_j \lambda_j x_{ij} + s_i^-.$

If $i \; \varepsilon \; I'$ then $x_{io} = 0$ so that $s_i^- = 0$ and for each j either $\lambda_j = 0$ or $x_{ij} = 0$. Consider some $j > k$. Then by (82) there will be some $i \; \varepsilon \; I'$ for which $x_{ij} \neq 0$ and, therefore, $\lambda_j = 0$ $(j = k + 1, \ldots, n)$. This means that the primal system for DMU_o in D reduces to that in D^*. Now, since by Theorems 6, 7, 8 and the arguments of Section 11 the category of DMU_o is determined by the primal system, the theorem is established.

The next corollaries all follow in turn from Theorem 18.

Corollary 19. There is at least one $j \leq k$ for which DMU_j belongs to E.

Corollary 20. Suppose that for some DMU_o the intersections $I_o \cap I_j$ are proper subsets of I_o for all $j \neq o$. Then DMU_o belongs to E.

36

Corollary 21. Suppose that for some DMU_o and some i we have $x_{io} = 0$ and $x_{ij} \neq 0$ for all $j \neq o$. Then DMU_o belongs to E.

Theorem 18 and its corollaries shed light on the importance and roles of zeros in the input coefficients x_{ij}. Any zero input coefficient will lead the way to an element of E.

A tempting early thought in developing DEA would be to assign the value "∞" to any formal quotient $f_j(u)/g_j(v)$ with numerator positive and denominator zero and then to call DMU_j scale efficient. The example

(84) $Y = [1\ 1\ 1]$, $X = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 2 & 0 \end{bmatrix}$

with n = 3 shows the fallacy of this approach. The virtual multiplier w with $u_1 = 1$, $v_1 = 1$, $v_2 = 0$ gives both $h_1(w)$ and $h_2(w)$ the form +/0. However, from the equations

$$y_2 = 1y_1 + 0y_2 + 0y_3$$
$$\frac{1}{2}x_2 = 1x_1 + 0x_2 + 0x_3$$

We see that for "0" = "2", the optimal θ^* is less than or equal to $\frac{1}{2}$ and hence DMU_2 is not scale efficient. (For this example $E = \{DMU_1, DMU_3\}$).

One might ask why we couldn't avoid the necessity for considering D-improper multipliers w of the type $f_j(u) = 0$ for all j by deleting any output r for which all $y_{rj} = 0$, i.e., forbid zero rows in Y? One reason is that in applying Theorem 18 we might obtain a reduced domain D^* for which both X^* and Y^* had zero rows even though this was not true for the X and Y of the parent domain D. The following example illustrates this point. Let D be defined by a matrix of the form

(85) $\begin{bmatrix} Y \\ -X \end{bmatrix} = \begin{bmatrix} Y^1 \\ 0 \\ -X^1 \\ 0 \end{bmatrix} \begin{bmatrix} 0_2 \\ Y^2 \\ 0_2 \\ -X^2 \end{bmatrix}$

where Y^1, Y^2, X^1, X^2 are matrices of appropriate degrees and without zero entries. Then both

$$D_1^* = \begin{bmatrix} Y^1 \\ 0_1 \\ -X^1 \\ 0 \end{bmatrix} \quad \text{and} \quad D_2^* = \begin{bmatrix} 0_2 \\ Y^2 \\ 0_2 \\ -X^2 \end{bmatrix}$$

can arise in applying Theorem 18. We arrive at D by combining the two data domains D_1 and D_2 defined by

$$D_1 = \begin{bmatrix} Y^1 \\ -X^1 \end{bmatrix} \quad , \quad D_2 = \begin{bmatrix} Y^2 \\ -X^2 \end{bmatrix}$$

and which have no common outputs or inputs.

The following example is more realistic

$$(86) \qquad \begin{bmatrix} Y \\ -X \end{bmatrix} = \begin{bmatrix} Y^1 & 0 & Y^{1,3} \\ 0 & Y^2 & Y^{2,3} \\ Y^{*,1} & Y^{*,2} & Y^{,3} \\ -X^1 & 0 & -X^{1,3} \\ 0 & -X^2 & -X^{2,3} \\ -X^{*,1} & -X^{*,2} & -X^{,3} \end{bmatrix}$$

and illustrates how permitting zero input and output coefficients extends the theory to situations where the DMUs do not have the same active inputs and outputs.

Some idea of the importance of being in a position to treat these zero inputs and/or outputs is apparent by reference to issues in divestiture such as were involved in the recent dissolution of AT&T. See [14]. In such cases, the issue turns on the complete removal of the entity from certain activities so that as in the AT&T case, zero outputs appear for local calls in one part of the proposed new entity and zero outputs appear for the toll call

components for other parts of the new entities arising from the divestitur

****

13. <u>THEORETICAL APPLICATION OF THE CLASSIFICATION THEOREMS.</u>

We turn now to methods for application of the classification theorems to a reduced data domain D. and consider three problems.

Problem 1. Partition the set J of DMU's into the six classes E, E', F, NE, NE', NF.

Problem 2. Identify all inclusion relations among the multiplier sets $W_1, \ldots, W_m$.

Problem 3. For each DMU_j construct optimal primal and dual vectors λ^j and w^j which satisfy the strong complimentary conditions (48) and (49).

We proceed as follows. First, (Phase 1) consider Problems 1 and 2 under the assumption that Problem 3 has been solved. Next, (Phase 2) see how to handle Problems 1 and 2 without having optimal, strongly complimentary solutions available. Then (Phase 3) turn to Problem 3.

<u>Phase 1</u>

Let λ^o, w^o be a strong complementary optimal pair for a selected DMU_o and let $\theta^* = z^* = \Sigma_r y_{ro} u_r^o$ be the optimal values for the objective functions (cf. Table 1).

Then it follows from the classification theorem that DMU_o belongs to:

E iff $\quad \theta^* = 1$, $s^{\neq o} = 0$, and $t_j^o > 0$ for all $j \neq 0$

E' iff $\quad \theta^* = 1$, $s^{\neq o} = 0$, and $t_j^o = 0$ for some $j \neq 0$

F iff $\quad \theta^* = 1$, $s^{\neq o} \neq 0$

NE iff $\quad \theta^* < 1$, $s^{\neq o} = 0$, and for some $DMUo'$

$\qquad\qquad \lambda_{o'}^o \neq 0$ and $\lambda_j^o = 0$ for all $j \neq o'$

39

NE' iff $\theta^* < 1$, $s^{\neq 0} = 0$, and more than one λ_j^0 is non-zero

NF iff $\theta^* < 1$, $s^{\neq 0} \neq 0$.

We observe that DMU_o is in NE iff there exists $\alpha > 0$ such that

$$\begin{bmatrix} y^o \\ -x^o \end{bmatrix} = \begin{bmatrix} \alpha\theta & y^{o'} \\ -\alpha x^{o'} \end{bmatrix} \text{ where } DMU_{o'} \in E \text{ the stringency of this condi-}$$

tion makes it likely that NE will be empty.

We complete Phase 1 by noting that $W_o \subset W_j$ iff $\lambda_j^0 \neq 0$.
For, if $\lambda_j^0 \neq 0$, $t_j = -\Sigma y_j u_r + \Sigma x_{ij} v_i = 0$ for all optimal

virtual multipliers $w = (u,v) \in W_o^m$ and therefore $W \in W_j$ (not

necessarily to W_j^m). It follows that $\lambda_j^0 \neq 0$ implies that

$W_o \subset W_j$. Conversely, if $\lambda_j^0 = 0$ then $t_j^0 \neq 0$. so that $w^o \notin W_j$ and

hence $W_o \not\subset W_j$.

Phase 2.

We begin with any optimal primal, dual pair (λ^*, w^*) for DMU_o.
Suppose that DMU_o belongs to RE, i.e., $\theta^* = z^* = 1$. Now consider
the modified primal program

$$\max \quad \Sigma_{r=1}^{s} s_r^+ + \Sigma_{i=1}^{m} s_i^-$$

subject to

$$\lambda_j \geq 0, \; s_r^+ \geq 0, \; s_i^- \geq 0 \quad (j = 1, \ldots, m, \; r = 1, \ldots s,$$

$$i = 1, \ldots, m)$$

and

$$P_o = \Sigma_{j=1}^{m} \lambda_j P_j + s^{\neq}.$$

Since $\theta^* = 1$, λ^* is a feasible solution of the modified system. If
the maximum is positive $DMU_o \in F$, otherwise $DMU_o \in EUE^1$.

40

In this case, we again solve a modified form of the primal problem; this time with the objective function min λ_o, but again with θ fixed at $\theta = 1$. If the minimum value is 1 then DMU_o belongs to E; otherwise (i.e., if the minimum is less than one) it belongs to E'.

If $DMU_o \in N$ (i.e., $\theta^* < 1$) then we apply this same procedure to the related $DMU_{o'}$ as defined in Section II above to determine which class NE, NE', or NF, contains DMU_o.

One minor point needs to be mentioned. If, perchance, $P_{o'}$ is proportional to some P_j for $DMU_j \in RE$ then the data domain $D \cup DMU_{o'}$ is not reduced. But, then DMU_o belongs to IE, IE' or EF according as DMU_j belongs to E, E', or F.

Problem 2 can be handled similarly. If $\lambda_j^* \neq 0$ then $W_o \subset W_j$. If $t_j^* \neq 0$ then $W_o \not\subset W_j$. If $\lambda_j^* = t_j^* = 0$ modify the primal program as follows. Fix θ at θ^* and change the objective function from "min θ" to "max λ_j". Then $W_o \subset W_j$ iff the optimal value of λ_j is positive. This completes Phase 2.

Phase 3

The strong complementary slackness theorem appeared in 1961 in the Charnes - Cooper text [7] pp. 441-443 where it is called the "Extended Theorem of the Alternative." Two years later Dantzig [17] gave the result in its slightly weaker homogenous form. In 1968 Balinski [1] Th. 8 p. 59, gave an elegant proof of the theorem under the title "Full Complementary Slackness." The alternate title used here "Strong Complementary Slackness" was introduced in 1970 in [18] Th. 1, p. 308. Since this theorem has remained relatively unnoticed by other writers we provide a brief discussion of a theoretical (not necessarily practical) computational procedure for obtaining λ^o, w^o using a standard simplex algorithm. We note that,

41

in general, we cannot expect λ^o, w^o to be basic solutions. However, we know that every optimal solution is a convex combination of basic solutions and we base our procedure on that property of λ^o and w^o.

Suppose now that $DMU_o \in RE$ and that λ^*, w^* are basic optimal solutions of the primal, dual systems with corresponding slack vectors

$$s^{\neq *} = \begin{bmatrix} s^{-*} \\ s^{+*} \end{bmatrix} \quad , \quad t^*.$$

Then, since $\theta^* = 1$, any solution $(\lambda, s^{\neq})$ of the modified primal constraint system

$$(87) \qquad P_o = P\lambda + s^{\neq}, \quad \lambda \geq 0, \quad s^{\neq} \geq 0$$

will be optimal for DMU_o.

Similarly, since $z^* = 1$, any solution (w, t) of the modified dual constraint sytem

$$(88) \qquad t + P^T w = 0, \quad t_o = 0, \quad x_o^T v = 1, \quad w \geq 0, \quad t \geq 0$$

will be optimal for DMU_o.

Now suppose for some i that $v_i^* = s_i^{-*} = 0$. Then exactly one of the problems

$$(89) \qquad \max s_i^- \text{ subject to } (87)$$

or

$$(90) \qquad \max v_i \text{ subject to } (88)$$

will have a positive objective value; denote a corresponding optimizing vector by $\lambda(i)$, $s^{\neq}(i)$ if s_i^- can be positive and by $v(i)$, $t(i)$ if v_i can be positive. Let $I^* = \left\{ i \mid v_i^* = s_i^{-*} = 0, \; v_i \text{ can be positive} \right\}$ and let $I^{**} = I - I^*$.

Next, let $v(I)$, $t(I)$ be the average of all pairs $v(i)$, $t(i)$ for $i \in I^*$ and let (I), $s(I)$ be the average of all pairs $\lambda(i)$, $s^{\neq}(i)$ for $i \in I^{**}$.

Then do the same for R and J to get $v(R)$, $t(R)$, $\lambda(R)$, $s(R)$ and $v(J)$, $t(J)$, $\lambda(J)$, $s(J)$. Now, the choice

$$(91) \quad (\lambda^0, s^0) = \frac{1}{4} [(\lambda^*, s^*) + (\lambda(I), s(I) + (\lambda(R), s(R)) + (\lambda(J), s(J))]$$

and

$$(92) \quad (w^0, t^0) = \frac{1}{4} [(w^*, t^*) + w(I), t(I) + (w(R), t(R) + (w(J), t(J)]$$

will give optimal primal and dual solutions which satisfy (49), i.e., which satisfy the strong complementary slackness condition.

We have now solved the three problems stated at the opening of this section. These solutions involve repeated use of simplex solutions to the dual problems in Table 1 plus closely related modifications. Although each such problem is simple, the number of the problems that needs to be addressed is very large. This leads to the problem of developing a more practical approach for securing numerical results in actual applications.

By an alternate formulation which is slightly more complex, one can arrive at an instant characterization of a DMU as an EUE', F or N. This new formulation suggests a strategy that may lead to more rapid and efficient solution procedures.

43

In preceding sections, the characterizations of E, E' and F are made in terms of the cone W and its cone subsets W_j. For evaluation of any particular DMU, designated as DMU_o, connection is made to the dual linear programming problems:

$$\min \theta$$

$$\max u^T y_o$$

(93)   subject to   subject to

$$Y\lambda \geq Y_o$$

$$u^T Y - v^T X \leq 0$$

$$\theta x_o - X\lambda \geq 0$$

$$v^T X_o = 1$$

$$\lambda \geq 0$$

$$u^T \geq 0, \ v^T \geq 0$$

But E, E' are characterized by W_o having a strictly positive vector set of dimension $m = s$ or less than $m + m$, respectively, whereas F has no strictly positive vector. These are not elements which are immediately related to the extreme points in (93) and thereby not to the computational methods e.g., simplex, etc. which rest on extreme points in (93).

Another linear programming problem which is extreme point computable and can distinguish F from $E \cup E'$ is:

Primal   Dual

$$\min \ -\theta^+ + \theta^-$$ $$\max \ \eta$$

           subject to   subject to

$$-y_o\theta^+ \quad -\psi^+ + Y\lambda \qquad \geq 0$$ $$-u^T \qquad \leq 0$$

$$x_o\theta^- \ -X\lambda \ -\psi^- \geq 0$$ $$-v_T \leq 0$$

(94) $$e^T\psi^+ \ + e^T\psi^- = 1$$ η otherwise unconstrained

$$\psi^+ \geq 0$$ $$\eta e^T - u^T \qquad \leq 0$$

$$\psi^- \geq 0$$ $$\eta e^T \qquad -v^T \leq 0$$

$$\lambda \ \geq 0$$ $$u^TY - v^TX \leq 0$$

$$\theta^+ \geq 0$$ $$-u^Ty_o \ \leq -1$$

$$\theta^- \geq 0$$ $$v^Tx_o \leq 1$$

Note the maximization problem is feasible for any scale efficient
DMU_o and is not feasible otherwise. Further, any feasible solution
must have $v^Tx_o = 1$. For if $v^Tx_o < 1$, then $u^Ty_o \leq v^Tx_o < 1$, and
infeasibility. The minimization problem is always feasible, however,
e.g., $\psi^+ = 0$, $\psi^- = x^o/e^Tx_o$, $\lambda_o = \theta^+$, $\lambda_j = 0$ for $j \neq 0$, $\theta^- = \theta^+ + 1/e^Tx_o$
satisfies the constraints.

As in past work by Charnes, Cooper and Seiford in non-linear
programming [12], we can distinguish the character of DMU_o by an
optimal solution (designated with asterisks) as follows

Theorem 19. DMU_o is $\left\{\begin{array}{l}E \text{ or } E' \\ F \\ \end{array}\right\}$ iff $-\theta^{+*} + \theta^{-*}$ is $\left\{\begin{array}{l} > 0 \\ = 0 \\ -\infty \end{array}\right\}$

Proof: Consider the maximal problem. When consistent, DMU_o is
scale efficient. Clearly, $w^* = (u^*, v^*) > 0$ iff $\eta^* = \min(u_r^*, V_i^*) > 0$.
But $\eta^* = -\theta^{+*} + \theta^{-*}$. Hence the first statement.

45

Next, note that η^* is the <u>maximum</u> of the minimal value of

the components of a feasible w for scale efficient DMU_o. Then

$0 = \eta^* - \theta^{+*} + \theta^{-*}$ means every w^* has at least one zero component,

or, DMU_o is in F.

Finally, since the minimal problem is always feasible, by the

duality states theorem for linear programming (see p. 486, Charnes,

Cooper, Duffuaa, Kress, [12]) the maximal problem is infeasible iff

the minimal problem has $-\theta^{+*} + \theta^{+*} = -\infty$.

****

14. CONCLUDING REMARKS

This completes what was to be done in the present paper. We

have now provided a general theoretical basis and characterization

of DEA solution sets which makes it possible to handle zero entries

in the data and in the variables. We have also provided a basis for

developing practical numerical algorithms with which to give these

ideas and concepts operational form for use in actual applications.

We have indicated some of the widened class of applications that

this makes possible for DEA.

REFERENCES

[1] Balinski, M.L., "Note on a Constructive Approach to Linear Programming," pp. 38-64 in Mathematics of the Decision Sciences, Part 1. Edited by G. B. Dantzig and A. F. Veinott, Jr. Lectures in Applied Mathematics, Vol. II (Providence, RI, American Mathematical Society, 1968).

[2] Banker, R. D., "Estimating Most Productive Scale Size Using Data Envelopment Analysis," European Journal of Operational Research 5, 1980, pp. 35-44..

[3] _____, A. Charnes and W. W. Cooper, "Models for Estimating Technical and Scale Efficiencies in Data Envelopment Analysis," Management Science 30, pp. 1078-1092.

[4] _____, _____ and _____, and A. Maindiratta, "A Comparison of DEA and Translog Estimates of Production Frontiers Using Simulated Observations from a Known Technology," in A. Dogramaci and R. Färe, Applications of Modern Production Theory: Efficiency and Productivity (Kluver, 1986).

[5] Bessent, A., W. Bessent, A. Charnes, W. W. Cooper and N. Thorogood, "Evaluation of Educational Proposals by Means of DEA," Educational Administration Quarterly 19, 1983, pp. 82-87.

[6] Bowlin, W. F., A. Charnes and H. D. Sherman, "Data Envelopment Analysis and Regression Approaches to Efficiency Evaluation and Estimation," Annals of Operations Research, 1985, pp. 113-138.

[7] Charnes, A. and W. W. Cooper, Management Models and Industrial Applications of Linear Programming, (New York: John WIley & Sons, Inc., 1961).

[8] _____ and _____, "Preface to Topics in Data Envelopment Analysis," Annals of Operations Research 2, 1985, pp. 59-94.

[9] _____ and _____, B. Golany, L. Seiford and J. Stutz, "Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions," Journal of Econometrics 30, 1985, pp. 91-107.

[10] _____ and _____, E. Rhodes, "Management Science Relations for Evaluation and Management Accountability," Journal of Enterprise Management 2, 1978, pp. 143-153.

[11] _____ and _____, _____, "Measuring Efficiency of Decision Making Units," European Journal of Operational Research 3, 1979, pp. 339.

[12] _____ and _____, S. Duffuaa, M. Kress, "Complexity and Computability of Solutions to Linear Programming Systems," International Journal of Computer and Information Sciences, 1980, pp. 483-506.

[13] _____, _____, and A. Schinnar, "Transforms and Approximations in Cost and Production Function Theory," _Omega_ 10, 1982, pp. 207-211.

[14] _____, _____, and T. Sueyoshi, "A Goal Programming/Constrained Regression Review of the Bell System Breakup," _Management Science_ (to appear).

[15] _____, _____, and _____, "A Goal Programming/Constrained Regression Study of AT&T as a Natural Monopoly," in O. A. Davis, ed. _Cost/Benefit Analysis_ (to appear).

[16] _____, _____, and R. M. Thrall, "Identifying and Classifying Scale and Technical Efficiencies in Data Envelopment Analysis," _Operations Research Letters_ (forthcoming).

[17] Dantzig, G. G., _Linear Programming and Extensions_ (Princeton: Princeton University Press, 1963).

[18] Spivey, W. Allen and Robert M. Thrall, _Linear Optimization_ (New York: Holt, Rinehart and Winston, IL, 1970).

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER CCS 512	2. GOVT ACCESSION NO. AD-AD6141	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* A Structure for Classifying and Characterizing Efficiencies and Inefficiencies in Data Envelopment Analysis		5. TYPE OF REPORT & PERIOD COVERED
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) A. Charnes, W.W. Cooper, R.M. Thrall		8. CONTRACT OR GRANT NUMBER(s) Noo14-86-C-0398
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Cybernetic Studies The University of Texas at Austin Austin, TX 78712		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research (Code 434) Washington, DC		12. REPORT DATE
		13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* Unclassified
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

This document has been approved for public release and sale; its distribution is unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

dual linear programs, multicriterion efficiency analysis, scale efficiency, strong complementary slackness, technical efficiency, virtual multipliers

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

DEA (Date Envelopment Analysis) attempts to identify sources and estimate amounts of inefficiencies contained in the outputs and inputs generated by managed entities called DMUs (= Decision Making Units). Explicit formulation of underlying functional relatons with specified parametric forms relating inputs to outputs is not required. An overall (scalar) measure of efficiency is also obtained for each DMU from the observed magnitudes of its multiple inputs and outputs without requiring use of a

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE
1 JAN 73
S/N 0102-014-6601 |

priori weights or relative value assumptions. A partition of DMUs into
six classes is established via primal and dual representation theorems
and three classification theorems which do not depend on non-archimedian
analysis. Earlier theory is extended to explain the consequences of
zero inputs and outputs and to utilize zero virtual multipliers (shadow
prices). Three of the six classes are scale inefficient and two of the
three scale efficient classes are also technically (= zero waste) efficient